DOCUMENT RESUME

ED 064 623 CG 007 143

AUTHOR Young, Jon I.; And Others
TITLE The Effects of Review Techniques and Instance Presentation on Concept Learning Tasks.
PUB DATE 72
NOTE 44p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Behavioral Science Research; College Students; *Concept Formation; Concept Teaching; *Learning; Learning Activities; Learning Experience; *Learning Processes; Learning Theories; Review (Reexamination); *Thought Processes

ABSTRACT

This study of concept learning was intended to supply data concerning the interaction of different values of three variables with each other. These three variables which may influence concept learning are augmentive information, method of presentation, and instance difficulty. Three types of specific review (total, partial, and none) were used, two methods of presentation (simultaneous and sequential) and two levels of task difficulty (easy instances and both hard and easy instances) were employed. Nine hypotheses were generated to examine the effects of these variables. Subjects for this study were 169 undergraduate psychology students at Brigham Young University. They were randomly assigned to one of twelve experimental groups or a control group. The results indicated that all of the treatment produced some learning by the subjects. It was concluded that the most efficient means of promoting concept learning is to use both easy and hard instances with specific review. The method of presentation was found to be optional. It was also concluded that the three variables mentioned above do seem to affect concept learning. (Author/BW)

ED 064623

CG (

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

THE EFFECTS OF REVIEW TECHNIQUES AND INSTANCE PRESENTATION ON CONCEPT LEARNING TASKS

Jon I. Young Kay H. Smith M. David Merrill

Brigham Young University

In 1969, Markle & Tiemann proposed that concept learning research should involve examining error responses, as well as correct classification responses. They proposed that these error responses be: (1) overgeneralization, an error which occurs when the subject incorrectly classifies negative instances as positive instances; (2) undergeneralization, another error which occurs when the subject incorrectly classifies positive instances as negative; and (3) misconception, which occurs when the subject incorrectly classifies either positive or negative instances.

Merrill and Boutwell (1973) propose three types of variables which may influence concept learning, and this influence may be measured in terms of both correct responses and one or more error responses. These three variables are augmentive information, method of presentation, and instance difficulty.

Augmentive information is so called because it is additional information provided to facilitate learning. Prompting and review are two examples of augmentive information. The literature indicates that prompts, if used excessively, reduce learning in memory tasks (Anderson, 1971), but increases concept learning (Merrill and Tennyson, 1972). Although it may also be that the differences result because of the amount of augmentive information, Merrill (1963; 1965, 1971) found a general review to be beneficial in concept learning, and he equated review with prompting as a means of promoting learning.

CG 007 143

The second variable discussed by Merrill and Boutwell (1973), was the method of presentation. Tennyson, Woolley and Merrill (1971) used a method of simultaneously presenting two positive and two negative instances so that the student could compare them, and therefore attention could be focused on the relevant attributes by systematically varying the irrelevant attributes.

Presenting instances sequentially would eliminate this opportunity to compare the instances, and the effect would be the same as if the instances were not matched. Tennyson, et al., found that unmatched instances produced more overgeneralization errors.

The third variable suggested by Merrill and Boutwell deals with instance difficulty. Instance difficulty has generally been considered to involve the number of relevant and irrelevant attributes. Tennyson, et al., found that undergeneralization occurs when only easy instances are used.

Although all three of these variables have been investigated, no research is available investigating the possible interactions in concept learning. The purpose of this study was to supply data concerning the interaction of different values of the variables with each other. Thus three types of specific review (total, partial, and none) were used, two methods of presentation (simultaneous and sequential) and two levels of task difficulty (easy instances and both hard and easy instances) were employed. The following hypotheses were generated to examine the effects of these variables.

Hypothesis 1: Any group which receives some treatment will be able to make significantly more correct classification responses on the posttest than the control group which receives no treatment.

Hypothesis 2: The number of correct classification responses made by subjects who learned with total specific review will be signifi-

cantly different from the number of correct classification responses made by subjects who learned with partial review.

Hypothesis 3: The number of correct classification responses made by subjects who learned with partial specific review will be significantly different from the number of correct classification responses made by subjects who learned with no specific review.

Hypothesis 4: When total specific review is employed, the proportion of undergeneralization responses made by the subjects in any one group will not be significantly different from the proportion of undergeneralization responses made by subjects in any other group.

Hypothesis 5: When total specific review is employed, the proportion of overgeneralization responses made by the subjects in any one group will not be significantly different from the proportion of overgeneralization responses made by subjects in any other groups.

Hypothesis 6: When no specific review is given, the number of correct classification responses made by the subjects who learned with simultaneous presentation method will be significantly different from the number of correct classification responses made by the subjects who learned with the sequential presentation method.

Hypothesis 7: When no specific review is given, the proportion of overgeneralization responses made by the subjects who learned with the sequential presentation method will be significantly

different from the proportion of overgeneralization responses made by subjects who learned with the simultaneous presentation method.

Hypothesis 8: When no specific review is given, the number of correct classification responses made by the subjects who learned with both easy and difficult instances will be significantly different from the number of correct classification responses made by the subjects who learned with only easy instances.

Hypothesis 9: When no specific review is given, the proportion of undergeneralization responses made by the subjects who learned with easy instances will be significantly different from the proportion of undergeneralization responses made by the subjects who learned with both easy and difficult instances.

METHOD

Subjects

Subjects for this study were students at Brigham Young University. One-half of the subjects were undergraduate students in educational psychology, and the remaining subjects were drawn from an introductory psychology course. All subjects participated in this research as partial fulfillment for course requirements of participating in research outside of class. A total of 169 subjects were used in twelve experimental groups and one control group.

Task

The concept used in this research was trochaic meter. Trochaic meter

is a rhythmic unit used in English poetry. This rhythmic unit is defined as a stressed syllable followed by an unstressed syllable (e.g., dán cíng) (Brewer, 1918; Deutsch, 1957). Trochaic meter is only one of five types of rhythmic units used in English poetry. This form of meter is fairly common in English because the great majority of two syllable words are perfect trochees (Wood, 1940).

Independent Variables

Three independent variables were used in this research: (1) amount of augmentive information; (2) method of presentation; (3) instance difficulty.

Specific review was used as augmentive information. When total specific review was used the subject was presented with the instances and directed to decide whether each instance was negative or positive. Following his decision he was told whether the instance was positive or negative, the correct meter was illustrated, and an explanation was given stating why the instance was so classified.

Partial specific review followed the same procedure, but only one-half of the instances had the meter illustrated and were followed by an explanation. The instances reviewed under this condition were randomly selected.

The third condition was actually a lack of augmentive information. The subject was merely told whether the instances was positive or negative.

The method of presentation variable involved two modes, simultaneous and sequential. Simultaneous presentation involved presenting two positive and two negative instances on the same page. Each positive instance was matched to a negative instance so that they had the same or similar content, number of lines, author, and stressed or unstressed last syllable. The two

positive instances were selected so that they differed as much as possible on these irrelevant attributes.

Sequential presentation involved presenting four instances, but each instance was given on a separate page.

Two conditions of instance difficulty were used. In one condition only easy instances were involved. In the second condition one-half of the instances were easy and one-half were difficult.

The degree of difficulty was taken from Tennyson, et al., (1971). In that study a total of ninety-two positive and negative instances were submitted, along with a definition of trochaic meter, to 140 undergraduate psychology students. The instances identified by at least 60% of the students were considered easy instances, and those instances identified by less than 49% of the students were considered difficult, or hard instances.

Dependent Variables

Responses of the subjects were classified as correct classification, overgeneralization, undergeneralization, and misconception. In addition, proportions of each of the error scores were computed. A covariant score was taken as the number of correct responses on a twenty-item pretest. The pretest items were of average difficulty (50 percentile range).

Procedure

The subjects were randomly assigned to one of twelve experimental groups or a control group. Table 1 summarizes all the experimental treatments. All groups were given a pretest consisting of ten positive and ten negative instances. The experimental subjects were then given the appropriate treatment.

Those groups which received both easy and hard instances were given four easy positive instances and four easy negative instances, as well as four difficult positive instances and four difficult negative instances, for a total of eight positive and eight negative instances. These same instances were used with both methods of presentation and with the three conditions of augmentive information.

The groups which received only easy instances were given eight easy positive instances and eight easy negative instances. These instances remained the same in both methods of presentation and in the three conditions of augmentive information.

The treatments were followed by a thirty-item posttest consisting of twelve positive instances and eighteen negative instances. The control group followed the same procedure, except that they read a short article on education instead of receiving the instructional material.

RESULTS

Seven dependent variables were obtained from the posttest. Four of these dependent variables were sum scores--correct classification responses, overgeneralization errors, undergeneralization errors, and misconception errors. Three of the dependent variables were proportional scores--proportion of overgeneralization responses, proportion of undergeneralization responses, and proportion of misconception responses. All seven dependent variables were analyzed by analysis of covariance with the number of correct responses on the pretest serving as the covariant for each dependent variable. The means for the dependent scores and the pretest scores are presented in Table 2.

No statistical comparisons were made between the correct classifications of the control group and any treatment group because the differences were obviously supportive of this hypothesis. The mean score for the treatment group which made the fewest correct classification was 17.44 and the pretest score was 4.54. In contrast, the mean score for correct classification of the control group was 1.153 and the pretest score was .923. These scores were low because only one subject was willing to make any choices on the pretest or the posttest. The other subjects in the control group refused to make any choices because they felt it futile. All thirteen subjects in this group reported that the irrelevant task did not help them in any way, nor did taking the pretest help them learn the concept of trochaic meter. For this reason--unwillingness to respond--the error responses were equally low, with the mean error scores being as follows: overgeneralization = .462; undergeneralization = .154; misconception = .385.

All of these mean scores were the direct result of the single subject responding to the pretest and posttest. His scores were as follows: pretest = 12; correct classification = 15; overgeneralization errors = 6; undergeneralization errors = 2; misconception errors = 5. From the difference between the pretest score and the posttest correct classification score it is obvious that little if any knowledge was acquired between the pretest and the posttest.

Partial support was found for Hypothesis 2. The mean score of correct classification responses for total specific review was 22.544, and for partial specific review the mean correct classification score was 19.721. When the means were compared using a conservative a priori test (Tukey (A), Weiner, 1968) no significant difference was found. However, the means are in the pre-

dicted direction with total specific review increasing the number of correct classification responses.

Hypothesis 3 also received partial support. The mean scores on correct classification response with no specific review was 18.157. The difference between this mean score and the mean score with partial specific review (19.721) was not significantly different, but again the means are in the predicted direction. The difference between the mean score for no specific review and the mean score for total specific review was significantly different ($p < .05$), with total specific review significantly increasing the number of correct classification responses. It was this difference that produced the significant main effect in the analysis of covariance of the three specific review conditions ($F = 16.85$, $p < .001$).

Hypothesis 4 was supported. The mean scores of the proportion of undergeneralization errors for the four groups that received total specific review (.4443; .5442; .5133; .5390) were not significantly different in the analysis of covariance.

Hypothesis 5 must be rejected. The mean scores of the four groups on the proportion of overgeneralization scores (.3469; .2516; .1908; .2144) showed a difference between the groups which learned with a simultaneous presentation and the groups which learned with the sequential presentation. The main effect for the difference between the methods of presentation was significant ($F = 4.44$, $p < .05$) with the means being .2992 for simultaneous presentation and .2026 for the sequential presentation method.

Hypothesis 6 must be rejected. When no specific review was given, the mean number of correct classification responses made with the simultaneous presentation method (13.24) was not different from the mean number of correct

classification responses made by the subjects who learned with the sequential presentation method (13.11) in the analysis of covariance ($F = 0.0$, $p > .05$).

Hypothesis 7 must also be rejected. The effect of the method of presentation was not significant ($F = 0.10$, $p > .05$). The mean score for the simultaneous presentation mode was .3318, and the mean score for the sequential presentation mode was .3455. With no specific review there was no significant difference between the methods of presentation on the proportion of overgeneralization errors, but with total specific review a significantly greater proportion of overgeneralization errors occurred with the simultaneous presentation method (Hypothesis 5).

Hypothesis 8 is not supported because the mean number of correct classification responses with both easy and hard instances (18.230) is not significantly different (analysis of covariance, $F = 0.0$, $p > .05$) from the mean number of correct classification responses when only easy instances are used (18.231).

Hypothesis 9 received support from the data. The mean proportion of undergeneralization responses when only easy instances were used was .4390, and when both easy and difficult instances were used the mean proportion of undergeneralization responses was .3230. This difference is significant ($F = 6.42$) at the .05 level.

While testing the hypotheses some interesting incidental results appeared. The specific review had an opposite effect on the proportion of overgeneralization and undergeneralization errors. While total specific review significantly reduced the proportion of overgeneralization errors over no review (Means: Total Review = .2554; Partial Review = .2985; No Review = .3432), it significantly increased the proportion of undergeneralization

errors (Means: Total Review = .5094; Partial Review = .4348; No Review = .3760).

A significant interaction was found between instance difficulty and the method of presentation for both proportions of undergeneralization responses and the proportion of overgeneralization responses. (Overgeneralization: $F = 3.92$, $p < .05$; Undergeneralization:: $F = 4.68$, $p < .001$). These interactions occurred when no specific review was given. Figure 1 illustrates these interactions. The figure shows that the overgeneralization and undergeneralization errors are in direct opposition to each other. Instance difficulty apparently affected the proportion of undergeneralization and overgeneralization errors only when simultaneous presentation was used. When easy instances alone were used with simultaneous presentation, the proportion of undergeneralization errors increased, while the proportion of overgeneralization errors decreased. The opposite occurred when both easy and difficult instances were used. This same interaction was found even when specific review was included, indicating that the instance difficulty variable affected the proportion of overgeneralization and undergeneralization responses, irrespective of the amount of the review variable.

A third significant interaction was found between the instance difficulty and augmentive information on overgeneralization ($F = 3.42$, $p < .05$). Figure 2 illustrates this interaction. A higher proportion of overgeneralization errors was found when both easy and hard instances were presented without specific review. Partial specific review on easy only instances produced the next highest proportion of overgeneralization errors.

All the sum scores were affected by the augmentive information variable. Specific review reduced the number of error scores by increasing the number of

correct classification responses. This was the only variable that significantly ($F = 7.04$, $p < .001$) affected the number of misconception errors. Total specific review produced the fewest misconception errors as indicated by a mean score of 1.83. Partial specific review followed with a mean score of 2.79, and no review had a mean score of 3.49.

The instance difficulty significantly affected both the correct classification responses and the undergeneralization responses. When both easy and difficult instances were used, the mean number of correct classification responses (20.93) was significantly greater ($F = 6.51$, $p < .025$) than the mean number of correct classification responses when only easy instances were used (19.35). When only easy instances were used, the mean number of undergeneralization responses (4.53) was significantly different ($F = 39.51$, $p < .001$) from the mean number of undergeneralization responses when both easy and difficult instances were used (3.19).

DISCUSSION

As a result of this study some of the hypotheses were accepted and others were rejected. All of the treatments produced some learning by the subjects. The control group with its irrelevant task was unable to show any learning gain between the pretest and the posttest.

The number of correct classification responses does not depend upon the method of presenting the instances. Both simultaneous and sequential methods of presentation produced the same number of correct classification responses with or without specific review. However, this variable may be important if the concept is very complicated, or if the positive and negative

instances can be contrasted through one or more of the senses (visual, auditory, tactile, etc.) Further research is needed to estimate the usefulness of mode of presentation with concepts of this more complicated nature.

Additionally, the difficulty of the instances used in the instruction does not affect the number of correct classification responses. It is probable that even the use of difficult instances alone would not significantly reduce the number of correct classification responses, especially if specific review is employed. However, this relationship needs to be studied experimentally.

Overgeneralization errors are reduced by specific review, but when total specific review is coupled with simultaneous presentation, the proportion of overgeneralization errors increases. The implications of this finding are that review and simultaneous presentation will cause students to classify some negative instances as positive instances. When specific review is not given, the proportion of overgeneralization responses does not differ between the modes of presentation. The conclusion is that if it is more important for the student to make correct choices, specific review should be used. However, if it is more crucial that the proportion of overgeneralization be reduced, then no specific review should be given.

Undergeneralization was shown by Woolley (1971) to occur when easy instances are the only ones used in the instruction. This study supports his conclusion. This study also showed that sequential presentation method will eliminate the difference between easy instances alone, and both easy and hard instances when the proportion of undergeneralization responses is concerned. Easy instances, coupled with simultaneous presentation, increases the proportion of undergeneralization responses. The conclusion regarding

instance difficulty is that the use of both easy and hard instances is best for reducing undergeneralization responses.

Woolley (1971) also concluded that undergeneralization results increase with an increase in information given the subject. This would follow from the fact that easy instances do give the subject more information; it was easier for subjects to identify them.

However, augmentive information in the form of specific review reduces the number of undergeneralization errors. Evidently the information received by virtue of the easy instances is different from that information received from the specific review. The exact difference between these two sources of information may be embodied in the combination of relevant and irrelevant attributes of the easy instances. Further research is needed to pinpoint the precise differences and similarities.

In summary, the most efficient means of promoting concept learning is using both easy and hard instances with specific review. The method of presentation is optional.

Theoretical Implications

The results of this study permit some conclusions about Merrill and Boutwell's (1973) assumptions and indicate areas wherein additional research is necessary.

The first variable considered was augmentive information in the form of specific review. It was this variable that had the strongest impact on correct classification scores. Although total specific review reduced all types of errors, it is most effective on misconception errors, followed by overgeneralization errors, and finally undergeneralization errors.

Misconception errors are produced when the subject is unable to correctly classify easy instances. Easy instances do not teach the limitations of the concept, which are represented by the difficult instances, and they are much easier to identify because they lack the complexity of the difficult instances. The addition of specific review with the easy instances serves to call attention to the already obvious relevant attributes, thereby increasing the likelihood that they will be correctly identified.

Undergeneralization errors tend to occur when learning involves more difficult instances. Calling attention to the relevant attributes of easy instances does not help in identification of difficult instances. Further research needs to answer two questions. First, is the amount or type of information transmitted to the learner by easy instances different from the amount or type of information transmitted by the more difficult instances? One possible finding may reflect the combinations of relevant and irrelevant attributes in the two types of instances. It is possible that certain combinations of attributes focus attention more readily and accurately on the relevant attributes. This may be the case for easy instances. Second, should different amounts or types of specific review be used to teach the instances, depending upon the degree of difficulty? This type of research may indicate that the combination of attributes in easy instances can be generalized to organize specific review, which will make difficult instances easier by making the relevant attributes more apparent.

The second variable that Merrill and Boutwell (1973) discuss is instance difficulty. Undergeneralization errors are produced when only easy instances are used in learning. This result is found even when specific review is used. It may be that the two variables are in fact transmitting

different information, or that specific review does not augment the information already transmitted by the instances. Research needs to be conducted in this area to determine if, in fact, the information is different. Research could also be used to determine if learner strategies are different when learning a concept with instances of various degrees of difficulty.

The third variable considered in this study was the method of presentation. Contrary to earlier findings, this study showed that the method of presentation does not have a significant effect on concept learning. The prediction was that sequential presentation would increase the number of overgeneralization errors because the student would not be able to compare the instances. Tennyson, et al., (1971) concluded that when positive and negative instances were not matched (matching permits comparisons), overgeneralization would occur. Sequential presentation should have produced the same result by preventing matching of the positive and negative instances. This was not shown to be the case.

One possible explanation involves the fact that this concept had only one relevant attribute. The complexity of the task may influence the advantages or disadvantages of the method of presentation. If the task is complex (involves several relevant and/or irrelevant attributes) then simultaneous presentation, in which the irrelevant attributes are systematically eliminated, may be the method of presentation most conducive to learning the concept.

In conclusion, these three variables--augmentive information, instance difficulty, and method of presentation--do seem to affect concept learning. However, these variables do not seem to be totally independent. Further research needs to investigate the relationships of these variables in terms of the amount and type of information being presented by the instances and

specific review, what types of learner strategies are employed, and the degree of task complexity. This information could then be used to propose a learning model specifying different methods of presentation and types of specific review, or prompting, which correspond to the nature of the task and the instances.

REFERENCES

Anderson, R. C. Two rules for manipulating the sequence of stimuli in concept learning tasks. (Unpublished paper, 1971)

Markle, S. M. and Tiemann, P. W. Really Understanding Concepts. Champaign, Illinois: Stipes, 1969.

Merrill, M. D. Learning and retention effects of a model and a preview in teaching an imaginary science. Urbana, Illinois: University of Illinois, Traning Research Laboratory, Technical Report No. 3, USOE Educational Media, Title VII contract 2-20-003, November, 1963.

Merrill, M. C. Correction and review on successive parts in learning a hierarchical task. Journal of Educational Psychology, 1965, 56, 225-234.

Merrill, M. D. and Tennyson, R. D. Concept acquisition as a function of relationships between positive and negative instances. Final Report of USOE Small Contract O-H-014, September, 1971.

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. Educational Technology, 1971. (In Press)

Merrill, M. D. and Boutwell, R. C. Instructional Development: Methodology and Research. Review of Research in Education, Vol. 1, 1973.

Tennyson, R. D., Woolley, F. R. and Merrill, M. D. Exemplar and non-exemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1970. (In Press)

Woolley, F. R. Effects of the presence of concept definition, pre-training concept exemplars, and feedback on the instruction of infinite conjunctive concepts. Unpublished dissertation, Brigham Young University, 1971.

TABLE 1

Experimental Design

GROUP	TREATMENT
I	Easy & Hard instances, Total Specific Review, Simultaneous presentation
II	Easy Instances, Total Specific Review, Simultaneous Presentation
III	Easy & Hard Instances, Partial Specific Review, Simultaneous Presentation
IV	Easy Instances, Partial Specific Review, Simultaneous Presentation
V	Easy & Hard Instances, No Specific Review, Simultaneous Presentation
VI	Easy Instances, No Specific Review, Simultaneous Presentation
VII	Easy & Hard Instances, Total Specific Review, Sequential Presentation
VIII	Easy Instances, Total Specific Review, Sequential Presentation
IX	Easy & Hard Instances, Partial Specific Review, Sequential Presentation
X	Easy Instances, Partial Specific Review, Sequential Presentation
XI	Easy & Hard Instances, No Specific Review, Sequential Presentation
XII	Easy Instances, No Specific Review, Sequential Presentation
XIII	Irrelevant Task

TABLE 2

ADJUSTED MEANS AND STANDARD DEVIATIONS
FOR DEPENDENT VARIABLE RAW SCORES

GROUP	SCORE	PRETEST	CORRECT CLASSIFICATION	OVER GENERALIZATION	UNDER GENERALIZATION	MISCONCEPT
	Mean	.62	23.55	2.62	2.37	1.46
I	SD	1.64	3.34	2.30	1.08	1.69
	Mean	2.69	21.30	2.70	4.00	2.00
II	SD	5.11	4.46	2.40	1.41	2.77
	Mean	.69	20.78	3.00	3.14	3.08
III	SD	1.26	5.70	2.81	1.51	2.56
	Mean	.31	19.18	4.06	4.60	2.16
IV	SD	.72	3.69	2.77	.84	1.75
	Mean	1.15	17.68	5.18	3.22	3.92
V	SD	2.18	3.03	1.80	1.31	2.27
	Mean	3.38	18.80	3.04	4.85	3.31
VI	SD	4.81	3.74	2.77	.95	2.16
	Mean	5.23	24.18	1.11	3.02	1.69
VII	SD	7.51	1.64	.88	1.24	.99
	Mean	1.08	21.14	2.17	4.53	2.15
VIII	SD	2.97	2.09	2.26	1.01	1.66
	Mean	2.00	20.64	2.75	3.77	2.85
IX	SD	4.66	3.45	1.97	1.67	2.63
	Mean	2.69	18.22	4.09	4.62	3.08
X	SD	4.63	2.55	1.90	1.00	1.77
	Mean	5.54	18.78	4.12	3.64	3.46
XI	SD	5.05	4.89	2.32	1.33	2.53
	Mean	4.54	17.44	4.70	4.55	3.31
XII	SD	6.57	4.05	2.40	1.55	2.55

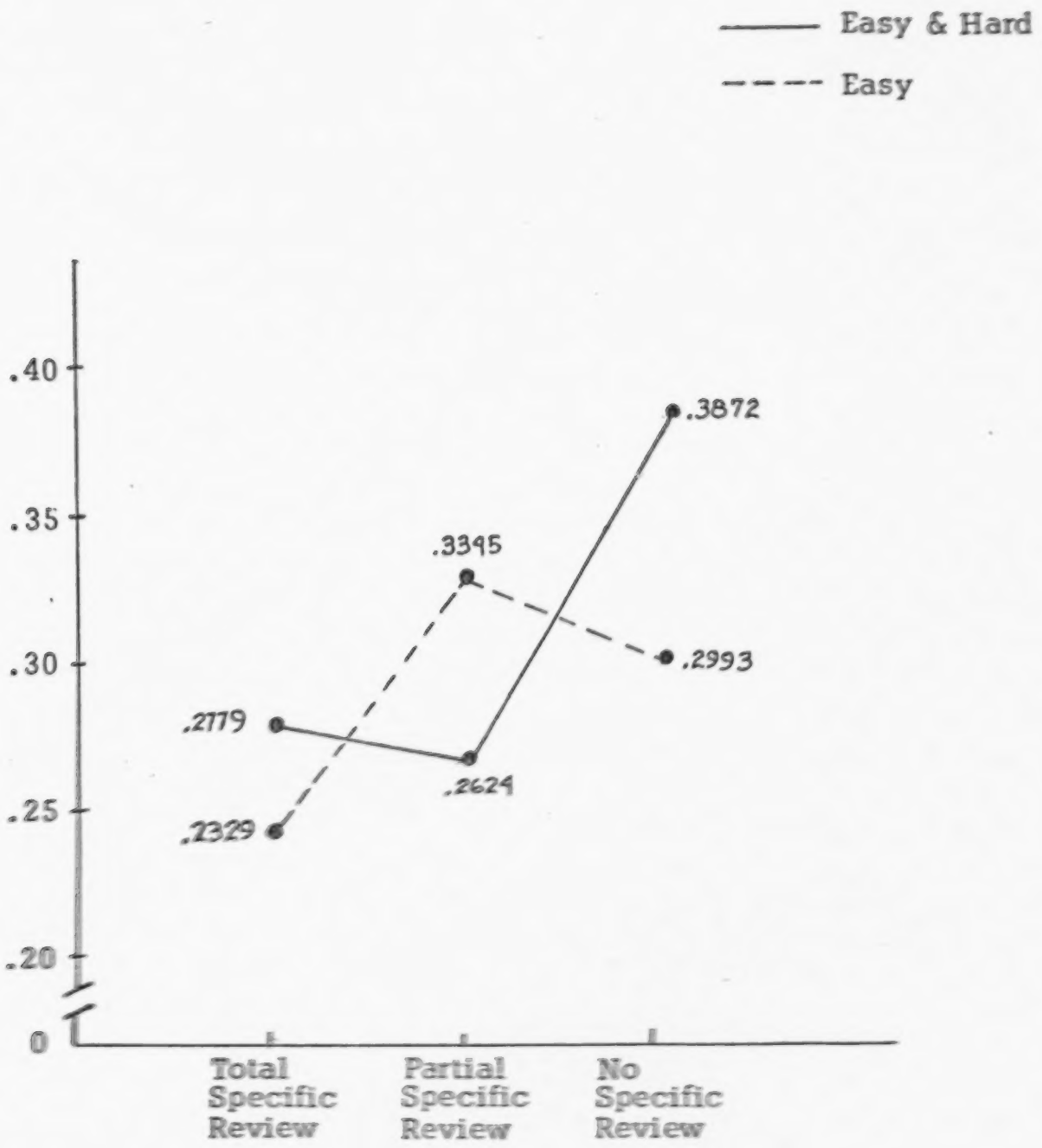

Figure 1. Proportion of Overgeneralization Errors for Specific Review and Instance Difficulty

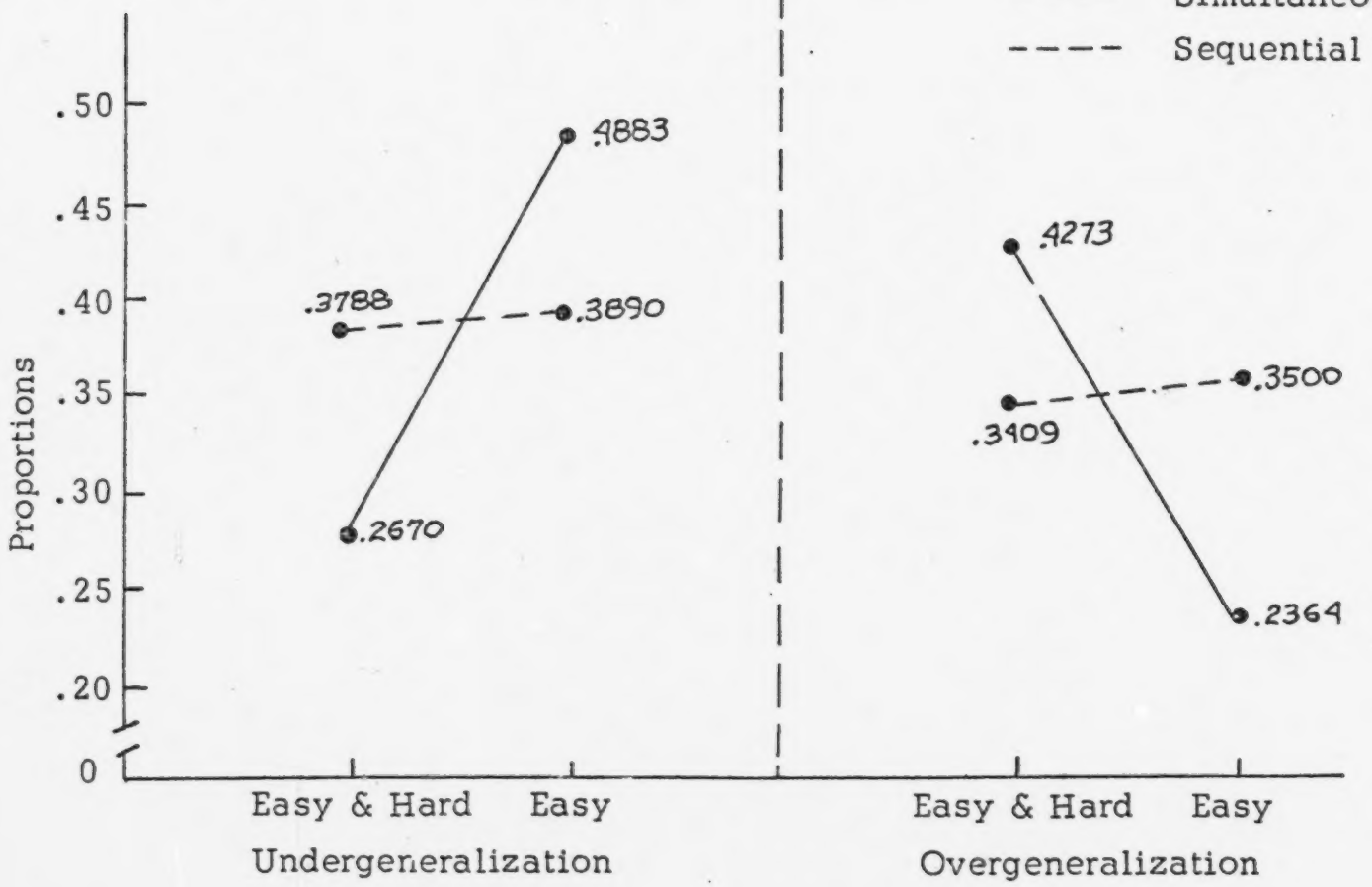

Figure 2. Interactions of Presentation and Instance Difficulty With No Specific Review On the Proportions of Undergeneralization and Overgeneralization Errors

THE EFFECTS OF REVIEW TECHNIQUES AND INSTANCE PRESENTATION ON CONCEPT LEARNING TASKS

Jon I. Young Kay H. Smith M. David Merrill

Brigham Young University

In 1969, Markle & Tiemann proposed that concept learning research should involve examining error responses, as well as correct classification responses. They proposed that these error responses be: (1) overgeneralization, an error which occurs when the subject incorrectly classifies negative instances as positive instances; (2) undergeneralization, another error which occurs when the subject incorrectly classifies positive instances as negative; and (3) misconception, which occurs when the subject incorrectly classifies either positive or negative instances.

Merrill and Boutwell (1973) propose three types of variables which may influence concept learning, and this influence may be measured in terms of both correct responses and one or more error responses. These three variables are augmentive information, method of presentation, and instance difficulty.

Augmentive information is so called because it is additional information provided to facilitate learning. Prompting and review are two examples of augmentive information. The literature indicates that prompts, if used excessively, reduce learning in memory tasks (Anderson, 1971), but increases concept learning (Merrill and Tennyson, 1972). Although it may also be that the differences result because of the amount of augmentive information, Merrill (1963; 1965, 1971) found a general review to be beneficial in concept learning, and he equated review with prompting as a means of promoting learning.

The second variable discussed by Merrill and Boutwell (1973), was the method of presentation. Tennyson, Woolley and Merrill (1971) used a method of simultaneously presenting two positive and two negative instances so that the student could compare them, and therefore attention could be focused on the relevant attributes by systematically varying the irrelevant attributes.

Presenting instances sequentially would eliminate this opportunity to compare the instances, and the effect would be the same as if the instances were not matched. Tennyson, et al., found that unmatched instances produced more overgeneralization errors.

The third variable suggested by Merrill and Boutwell deals with instance difficulty. Instance difficulty has generally been considered to involve the number of relevant and irrelevant attributes. Tennyson, et al., found that undergeneralization occurs when only easy instances are used.

Although all three of these variables have been investigated, no research is available investigating the possible interactions in concept learning. The purpose of this study was to supply data concerning the interaction of different values of the variables with each other. Thus three types of specific review (total, partial, and none) were used, two methods of presentation (simultaneous and sequential) and two levels of task difficulty (easy instances and both hard and easy instances) were employed. The following hypotheses were generated to examine the effects of these variables.

Hypothesis 1: Any group which receives some treatment will be able to make significantly more correct classification responses on the posttest than the control group which receives no treatment.

Hypothesis 2: The number of correct classification responses made by subjects who learned with total specific review will be signifi-

cantly different from the number of correct classification responses made by subjects who learned with partial review.

Hypothesis 3: The number of correct classification responses made by subjects who learned with partial specific review will be significantly different from the number of correct classification responses made by subjects who learned with no specific review.

Hypothesis 4: When total specific review is employed, the proportion of undergeneralization responses made by the subjects in any one group will not be significantly different from the proportion of undergeneralization responses made by subjects in any other group.

Hypothesis 5: When total specific review is employed, the proportion of overgeneralization responses made by the subjects in any one group will not be significantly different from the proportion of overgeneralization responses made by subjects in any other groups.

Hypothesis 6: When no specific review is given, the number of correct classification responses made by the subjects who learned with simultaneous presentation method will be significantly different from the number of correct classification responses made by the subjects who learned with the sequential presentation method.

Hypothesis 7: When no specific review is given, the proportion of overgeneralization responses made by the subjects who learned with the sequential presentation method will be significantly

different from the proportion of overgeneralization responses made by subjects who learned with the simultaneous presentation method.

Hypothesis 8: When no specific review is given, the number of correct classification responses made by the subjects who learned with both easy and difficult instances will be significantly different from the number of correct classification responses made by the subjects who learned with only easy instances.

Hypothesis 9: When no specific review is given, the proportion of undergeneralization responses made by the subjects who learned with easy instances will be significantly different from the proportion of undergeneralization responses made by the subjects who learned with both easy and difficult instances.

METHOD

Subjects

Subjects for this study were students at Brigham Young University. One-half of the subjects were undergraduate students in educational psychology, and the remaining subjects were drawn from an introductory psychology course. All subjects participated in this research as partial fulfillment for course requirements of participating in research outside of class. A total of 169 subjects were used in twelve experimental groups and one control group.

Task

The concept used in this research was trochaic meter. Trochaic meter

is a rhythmic unit used in English poetry. This rhythmic unit is defined as a stressed syllable followed by an unstressed syllable (e.g., dán cĭng) (Brewer, 1918; Deutsch, 1957). Trochaic meter is only one of five types of rhythmic units used in English poetry. This form of meter is fairly common in English because the great majority of two syllable words are perfect trochees (Wood, 1940).

Independent Variables

Three independent variables were used in this research: (1) amount of augmentive information; (2) method of presentation; (3) instance difficulty.

Specific review was used as augmentive information. When total specific review was used the subject was presented with the instances and directed to decide whether each instance was negative or positive. Following his decision he was told whether the instance was positive or negative, the correct meter was illustrated, and an explanation was given stating why the instance was so classified.

Partial specific review followed the same procedure, but only one-half of the instances had the meter illustrated and were followed by an explanation. The instances reviewed under this condition were randomly selected.

The third condition was actually a lack of augmentive information. The subject was merely told whether the instances was positive or negative.

The method of presentation variable involved two modes, simultaneous and sequential. Simultaneous presentation involved presenting two positive and two negative instances on the same page. Each positive instance was matched to a negative instance so that they had the same or similar content, number of lines, author, and stressed or unstressed last syllable. The two

positive instances were selected so that they differed as much as possible on these irrelevant attributes.

Sequential presentation involved presenting four instances, but each instance was given on a separate page.

Two conditions of instance difficulty were used. In one condition only easy instances were involved. In the second condition one-half of the instances were easy and one-half were difficult.

The degree of difficulty was taken from Tennyson, et al., (1971). In that study a total of ninety-two positive and negative instances were submitted, along with a definition of trochaic meter, to 140 undergraduate psychology students. The instances identified by at least 60% of the students were considered easy instances, and those instances identified by less than 49% of the students were considered difficult, or hard instances.

Dependent Variables

Responses of the subjects were classified as correct classification, overgeneralization, undergeneralization, and misconception. In addition, proportions of each of the error scores were computed. A covariant score was taken as the number of correct responses on a twenty-item pretest. The pretest items were of average difficulty (50 percentile range).

Procedure

The subjects were randomly assigned to one of twelve experimental groups or a control group. Table 1 summarizes all the experimental treatments. All groups were given a pretest consisting of ten positive and ten negative instances. The experimental subjects were then given the appropriate treatment.

Those groups which received both easy and hard instances were given four easy positive instances and four easy negative instances, as well as four difficult positive instances and four difficult negative instances, for a total of eight positive and eight negative instances. These same instances were used with both methods of presentation and with the three conditions of augmentive information.

The groups which received only easy instances were given eight easy positive instances and eight easy negative instances. These instances remained the same in both methods of presentation and in the three conditions of augmentive information.

The treatments were followed by a thirty-item posttest consisting of twelve positive instances and eighteen negative instances. The control group followed the same procedure, except that they read a short article on education instead of receiving the instructional material.

RESULTS

Seven dependent variables were obtained from the posttest. Four of these dependent variables were sum scores--correct classification responses, overgeneralization errors, undergeneralization errors, and misconception errors. Three of the dependent variables were proportional scores--proportion of overgeneralization responses, proportion of undergeneralization responses, and proportion of misconception responses. All seven dependent variables were analyzed by analysis of covariance with the number of correct responses on the pretest serving as the covariant for each dependent variable. The means for the dependent scores and the pretest scores are presented in Table 2.

No statistical comparisons were made between the correct classifications of the control group and any treatment group because the differences were obviously supportive of this hypothesis. The mean score for the treatment group which made the fewest correct classification was 17.44 and the pretest score was 4.54. In contrast, the mean score for correct classification of the control group was 1.153 and the pretest score was .923. These scores were low because only one subject was willing to make any choices on the pretest or the posttest. The other subjects in the control group refused to make any choices because they felt it futile. All thirteen subjects in this group reported that the irrelevant task did not help them in any way, nor did taking the pretest help them learn the concept of trochaic meter. For this reason--unwillingness to respond--the error responses were equally low, with the mean error scores being as follows: overgeneralization = .462; undergeneralization = .154; misconception = .385.

All of these mean scores were the direct result of the single subject responding to the pretest and posttest. His scores were as follows: pretest = 12; correct classification = 15; overgeneralization errors = 6; undergeneralization errors = 2; misconception errors = 5. From the difference between the pretest score and the posttest correct classification score it is obvious that little if any knowledge was acquired between the pretest and the posttest.

Partial support was found for Hypothesis 2. The mean score of correct classification responses for total specific review was 22.544, and for partial specific review the mean correct classification score was 19.721. When the means were compared using a conservative a priori test (Tukey (A), Weiner, 1968) no significant difference was found. However, the means are in the pre-

dicted direction with total specific review increasing the number of correct classification responses.

Hypothesis 3 also received partial support. The mean scores on correct classification response with no specific review was 18.157. The difference between this mean score and the mean score with partial specific review (19.721) was not significantly different, but again the means are in the predicted direction. The difference between the mean score for no specific review and the mean score for total specific review was significantly different ($p < .05$), with total specific review significantly increasing the number of correct classification responses. It was this difference that produced the significant main effect in the analysis of covariance of the three specific review conditions ($F = 16.85$, $p < .001$).

Hypothesis 4 was supported. The mean scores of the proportion of undergeneralization errors for the four groups that received total specific review (.4443; .5442; .5133; .5390) were not significantly different in the analysis of covariance.

Hypothesis 5 must be rejected. The mean scores of the four groups on the proportion of overgeneralization scores (.3469; .2516; .1908; .2144) showed a difference between the groups which learned with a simultaneous presentation and the groups which learned with the sequential presentation. The main effect for the difference between the methods of presentation was significant ($F = 4.44$, $p < .05$) with the means being .2992 for simultaneous presentation and .2026 for the sequential presentation method.

Hypothesis 6 must be rejected. When no specific review was given, the mean number of correct classification responses made with the simultaneous presentation method (13.24) was not different from the mean number of correct

classification responses made by the subjects who learned with the sequential presentation method (13.11) in the analysis of covariance ($F = 0.0$, $p > .05$).

Hypothesis 7 must also be rejected. The effect of the method of presentation was not significant ($F = 0.10$, $p > .05$). The mean score for the simultaneous presentation mode was .3318, and the mean score for the sequential presentation mode was .3455. With no specific review there was no significant difference between the methods of presentation on the proportion of overgeneralization errors, but with total specific review a significantly greater proportion of overgeneralization errors occurred with the simultaneous presentation method (Hypothesis 5).

Hypothesis 8 is not supported because the mean number of correct classification responses with both easy and hard instances (18.230) is not significantly different (analysis of covariance, $F = 0.0$, $p > .05$) from the mean number of correct classification responses when only easy instances are used (18.231).

Hypothesis 9 received support from the data. The mean proportion of undergeneralization responses when only easy instances were used was .4390, and when both easy and difficult instances were used the mean proportion of undergeneralization responses was .3230. This difference is significant ($F = 6.42$) at the .05 level.

While testing the hypotheses some interesting incidental results appeared. The specific review had an opposite effect on the proportion of overgeneralization and undergeneralization errors. While total specific review significantly reduced the proportion of overgeneralization errors over no review (Means: Total Review = .2554; Partial Review = .2985; No Review = .3432), it significantly increased the proportion of undergeneralization

errors (Means: Total Review = .5094; Partial Review = .4348; No Review = .3760).

A significant interaction was found between instance difficulty and the method of presentation for both proportions of undergeneralization responses and the proportion of overgeneralization responses. (Overgeneralization: $F = 3.92$, $p < .05$; Undergeneralization:: $F = 4.68$, $p < .001$). These interactions occurred when no specific review was given. Figure 1 illustrates these interactions. The figure shows that the overgeneralization and undergeneralization errors are in direct opposition to each other. Instance difficulty apparently affected the proportion of undergeneralization and overgeneralization errors only when simultaneous presentation was used. When easy instances alone were used with simultaneous presentation, the proportion of undergeneralization errors increased, while the proportion of overgeneralization errors decreased. The opposite occurred when both easy and difficult instances were used. This same interaction was found even when specific review was included, indicating that the instance difficulty variable affected the proportion of overgeneralization and undergeneralization responses, irrespective of the amount of the review variable.

A third significant interaction was found between the instance difficulty and augmentive information on overgeneralization ($F = 3.42$, $p < .05$). Figure 2 illustrates this interaction. A higher proportion of overgeneralization errors was found when both easy and hard instances were presented without specific review. Partial specific review on easy only instances produced the next highest proportion of overgeneralization errors.

All the sum scores were affected by the augmentive information variable. Specific review reduced the number of error scores by increasing the number of

correct classification responses. This was the only variable that significantly ($F = 7.04$, $p < .001$) affected the number of misconception errors. Total specific review produced the fewest misconception errors as indicated by a mean score of 1.83. Partial specific review followed with a mean score of 2.79, and no review had a mean score of 3.49.

The instance difficulty significantly affected both the correct classification responses and the undergeneralization responses. When both easy and difficult instances were used, the mean number of correct classification responses (20.93) was significantly greater ($F = 6.51$, $p < .025$) than the mean number of correct classification responses when only easy instances were used (19.35). When only easy instances were used, the mean number of undergeneralization responses (4.53) was significantly different ($F = 39.51$, $p < .001$) from the mean number of undergeneralization responses when both easy and difficult instances were used (3.19).

DISCUSSION

As a result of this study some of the hypotheses were accepted and others were rejected. All of the treatments produced some learning by the subjects. The control group with its irrelevant task was unable to show any learning gain between the pretest and the posttest.

The number of correct classification responses does not depend upon the method of presenting the instances. Both simultaneous and sequential methods of presentation produced the same number of correct classification responses with or without specific review. However, this variable may be important if the concept is very complicated, or if the positive and negative

instances can be contrasted through one or more of the senses (visual, auditory, tactile, etc.) Further research is needed to estimate the usefulness of mode of presentation with concepts of this more complicated nature.

Additionally, the difficulty of the instances used in the instruction does not affect the number of correct classification responses. It is probable that even the use of difficult instances alone would not significantly reduce the number of correct classification responses, especially if specific review is employed. However, this relationship needs to be studied experimentally.

Overgeneralization errors are reduced by specific review, but when total specific review is coupled with simultaneous presentation, the proportion of overgeneralization errors increases. The implications of this finding are that review and simultaneous presentation will cause students to classify some negative instances as positive instances. When specific review is not given, the proportion of overgeneralization responses does not differ between the modes of presentation. The conclusion is that if it is more important for the student to make correct choices, specific review should be used. However, if it is more crucial that the proportion of overgeneralization be reduced, then no specific review should be given.

Undergeneralization was shown by Woolley (1971) to occur when easy instances are the only ones used in the instruction. This study supports his conclusion. This study also showed that sequential presentation method will eliminate the difference between easy instances alone, and both easy and hard instances when the proportion of undergeneralization responses is concerned. Easy instances, coupled with simultaneous presentation, increases the proportion of undergeneralization responses. The conclusion regarding

instance difficulty is that the use of both easy and hard instances is best for reducing undergeneralization responses.

Woolley (1971) also concluded that undergeneralization results increase with an increase in information given the subject. This would follow from the fact that easy instances do give the subject more information; it was easier for subjects to identify them.

However, augmentive information in the form of specific review reduces the number of undergeneralization errors. Evidently the information received by virtue of the easy instances is different from that information received from the specific review. The exact difference between these two sources of information may be embodied in the combination of relevant and irrelevant attributes of the easy instances. Further research is needed to pinpoint the precise differences and similarities.

In summary, the most efficient means of promoting concept learning is using both easy and hard instances with specific review. The method of presentation is optional.

Theoretical Implications

The results of this study permit some conclusions about Merrill and Boutwell's (1973) assumptions and indicate areas wherein additional research is necessary.

The first variable considered was augmentive information in the form of specific review. It was this variable that had the strongest impact on correct classification scores. Although total specific review reduced all types of errors, it is most effective on misconception errors, followed by overgeneralization errors, and finally undergeneralization errors.

Misconception errors are produced when the subject is unable to correctly classify easy instances. Easy instances do not teach the limitations of the concept, which are represented by the difficult instances, and they are much easier to identify because they lack the complexity of the difficult instances. The addition of specific review with the easy instances serves to call attention to the already obvious relevant attributes, thereby increasing the likelihood that they will be correctly identified.

Undergeneralization errors tend to occur when learning involves more difficult instances. Calling attention to the relevant attributes of easy instances does not help in identification of difficult instances. Further research needs to answer two questions. First, is the amount or type of information transmitted to the learner by easy instances different from the amount or type of information transmitted by the more difficult instances? One possible finding may reflect the combinations of relevant and irrelevant attributes in the two types of instances. It is possible that certain combinations of attributes focus attention more readily and accurately on the relevant attributes. This may be the case for easy instances. Second, should different amounts or types of specific review be used to teach the instances, depending upon the degree of difficulty? This type of research may indicate that the combination of attributes in easy instances can be generalized to organize specific review, which will make difficult instances easier by making the relevant attributes more apparent.

The second variable that Merrill and Boutwell (1973) discuss is instance difficulty. Undergeneralization errors are produced when only easy instances are used in learning. This result is found even when specific review is used. It may be that the two variables are in fact transmitting

different information, or that specific review does not augment the information already transmitted by the instances. Research needs to be conducted in this area to determine if, in fact, the information is different. Research could also be used to determine if learner strategies are different when learning a concept with instances of various degrees of difficulty.

The third variable considered in this study was the method of presentation. Contrary to earlier findings, this study showed that the method of presentation does not have a significant effect on concept learning. The prediction was that sequential presentation would increase the number of overgeneralization errors because the student would not be able to compare the instances. Tennyson, et al., (1971) concluded that when positive and negative instances were not matched (matching permits comparisons), overgeneralization would occur. Sequential presentation should have produced the same result by preventing matching of the positive and negative instances. This was not shown to be the case.

One possible explanation involves the fact that this concept had only one relevant attribute. The complexity of the task may influence the advantages or disadvantages of the method of presentation. If the task is complex (involves several relevant and/or irrelevant attributes) then simultaneous presentation, in which the irrelevant attributes are systematically eliminated, may be the method of presentation most conducive to learning the concept.

In conclusion, these three variables--augmentive information, instance difficulty, and method of presentation--do seem to affect concept learning. However, these variables do not seem to be totally independent. Further research needs to investigate the relationships of these variables in terms of the amount and type of information being presented by the instances and

specific review, what types of learner strategies are employed, and the degree of task complexity. This information could then be used to propose a learning model specifying different methods of presentation and types of specific review, or prompting, which correspond to the nature of the task and the instances.

REFERENCES

Anderson, R. C. Two rules for manipulating the sequence of stimuli in concept learning tasks. (Unpublished paper, 1971)

Markle, S. M. and Tiemann, P. W. Really Understanding Concepts. Champaign, Illinois: Stipes, 1969.

Merrill, M. D. Learning and retention effects of a model and a preview in teaching an imaginary science. Urbana, Illinois: University of Illinois, Traning Research Laboratory, Technical Report No. 3, USOE Educational Media, Title VII contract 2-20-003, November, 1963.

Merrill, M. C. Correction and review on successive parts in learning a hierarchical task. Journal of Educational Psychology, 1965, 56, 225-234.

Merrill, M. D. and Tennyson, R. D. Concept acquisition as a function of relationships between positive and negative instances. Final Report of USOE Small Contract O-H-014, September, 1971.

Merrill, M. D. Necessary psychological conditions for defining instructional outcomes. Educational Technology, 1971. (In Press)

Merrill, M. D. and Boutwell, R. C. Instructional Development: Methodology and Research. Review of Research in Education, Vol. 1, 1973.

Tennyson, R. D., Woolley, F. R. and Merrill, M. D. Exemplar and non-exemplar variables which produce correct concept classification behavior and specified classification errors. Journal of Educational Psychology, 1970. (In Press)

Woolley, F. R. Effects of the presence of concept definition, pre-training concept exemplars, and feedback on the instruction of infinite conjunctive concepts. Unpublished dissertation, Brigham Young University, 1971.

TABLE 1

Experimental Design

GROUP	TREATMENT
I	Easy & Hard instances, Total Specific Review, Simultaneous presentation
II	Easy Instances, Total Specific Review, Simultaneous Presentation
III	Easy & Hard Instances, Partial Specific Review, Simultaneous Presentation
IV	Easy Instances, Partial Specific Review, Simultaneous Presentation
V	Easy & Hard Instances, No Specific Review, Simultaneous Presentation
VI	Easy Instances, No Specific Review, Simultaneous Presentation
VII	Easy & Hard Instances, Total Specific Review, Sequential Presentation
VIII	Easy Instances, Total Specific Review, Sequential Presentation
IX	Easy & Hard Instances, Partial Specific Review, Sequential Presentation
X	Easy Instances, Partial Specific Review, Sequential Presentation
XI	Easy & Hard Instances, No Specific Review, Sequential Presentation
XII	Easy Instances, No Specific Review, Sequential Presentation
XIII	Irrelevant Task

TABLE 2

ADJUSTED MEANS AND STANDARD DEVIATIONS
FOR DEPENDENT VARIABLE RAW SCORES

GROUP	SCORE	PRETEST	CORRECT CLASSIFICATION	OVER GENERALIZATION	UNDER GENERALIZATION	MISCONCEPT
I	Mean	.62	23.55	2.62	2.37	1.46
	SD	1.64	3.34	2.30	1.08	1.69
II	Mean	2.69	21.30	2.70	4.00	2.00
	SD	5.11	4.46	2.40	1.41	2.77
III	Mean	.69	20.78	3.00	3.14	3.08
	SD	1.26	5.70	2.81	1.51	2.56
IV	Mean	.31	19.18	4.06	4.60	2.16
	SD	.72	3.69	2.77	.84	1.75
V	Mean	1.15	17.68	5.18	3.22	3.92
	SD	2.18	3.03	1.80	1.31	2.27
VI	Mean	3.38	18.80	3.04	4.85	3.31
	SD	4.81	3.74	2.77	.95	2.16
VII	Mean	5.23	24.18	1.11	3.02	1.69
	SD	7.51	1.64	.88	1.24	.99
VIII	Mean	1.08	21.14	2.17	4.53	2.15
	SD	2.97	2.09	2.26	1.01	1.66
IX	Mean	2.00	20.64	2.75	3.77	2.85
	SD	4.66	3.45	1.97	1.67	2.63
X	Mean	2.69	18.22	4.09	4.62	3.08
	SD	4.63	2.55	1.90	1.00	1.77
XI	Mean	5.54	18.78	4.12	3.64	3.46
	SD	5.05	4.89	2.32	1.33	2.53
XII	Mean	4.54	17.44	4.70	4.55	3.31
	SD	6.57	4.05	2.40	1.55	2.55

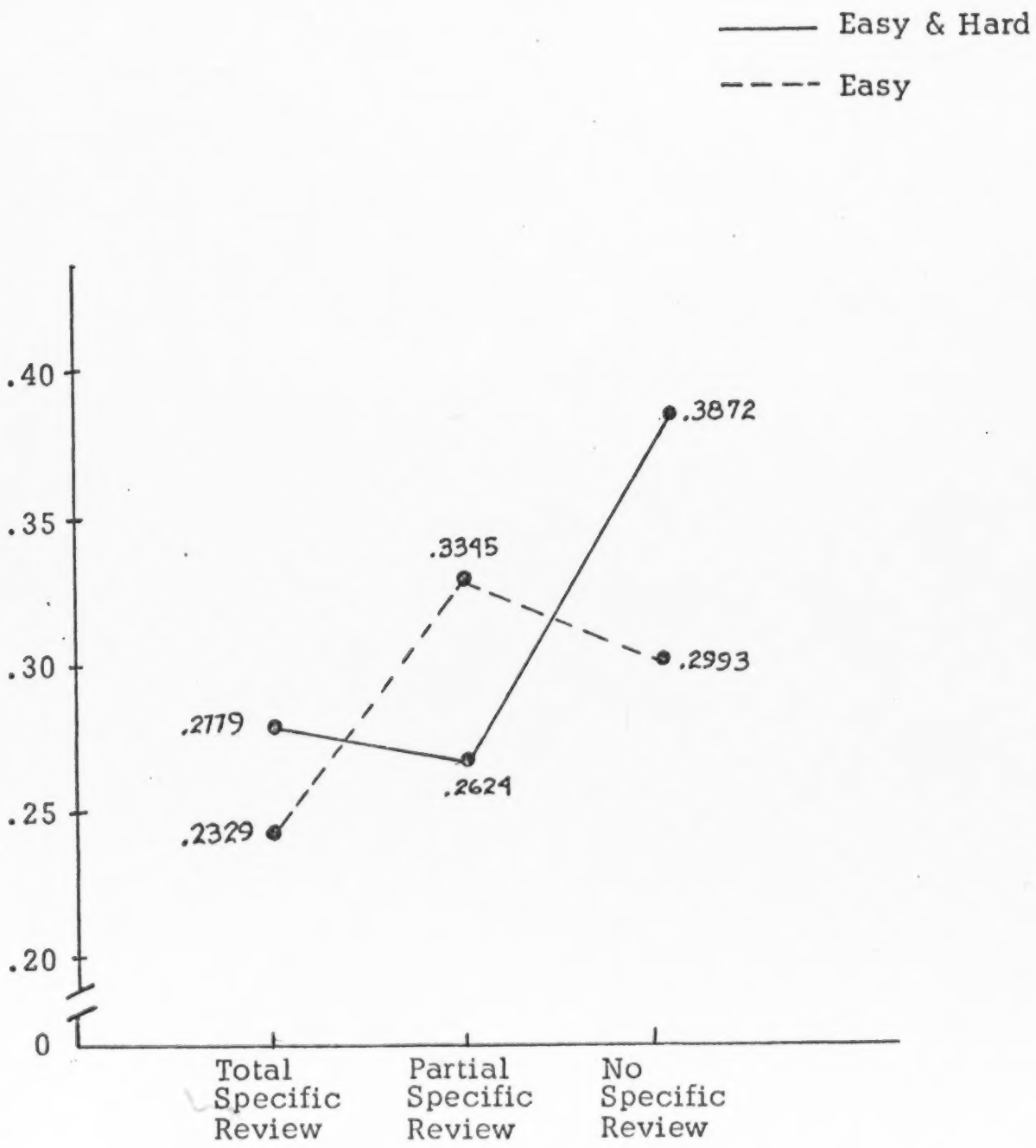

Figure 1. Proportion of Overgeneralization Errors for Specific Review and Instance Difficulty

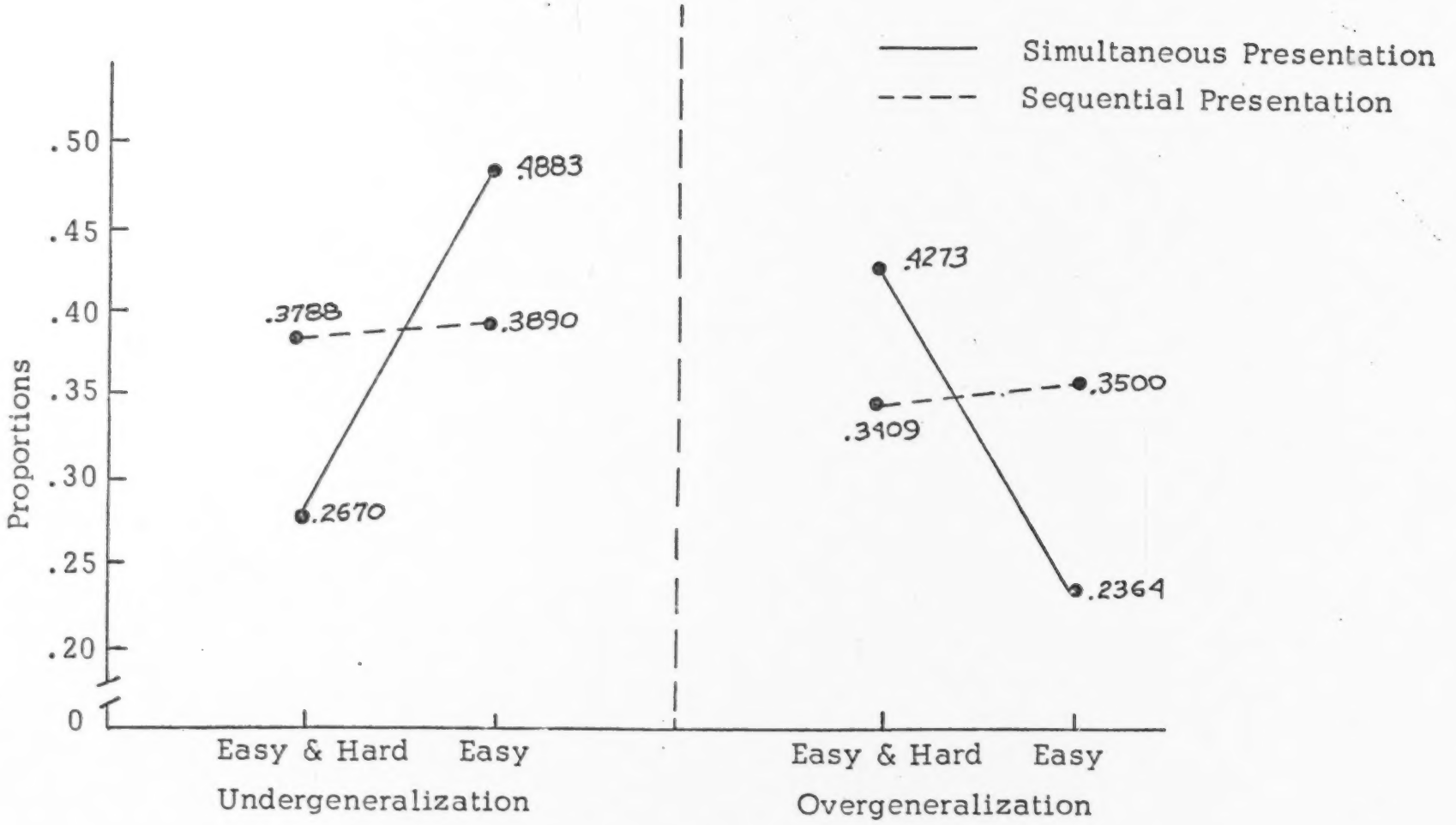

Figure 2. Interactions of Presentation and Instance Difficulty With No Specific Review On the Proportions of Undergeneralization and Overgeneralization Errors